Common Gateway Interface (CGI)

Introduction to CGI

Common Gateway Interface (CGI) is a standard for interfacing external programs with information servers on the Internet. So what does this mean? Basically, CGI is distinguished from a plain HTML document in that the plain HTML document is static, while CGI executes in real-time to output *dynamic* information. A program that implements CGI is executable, while the plain HTML document exists as a constant text file that doesn't change. CGI, then, obtains information from users and tailors pages to their needs. While there are newer ways to perform the same kinds of actions that traditionally have been implemented with CGI, the latter is older and, in many ways, more versatile. It is for this reason that, over time, CGI has become generalized to refer to any program that runs on a Web server and interacts with a browser.

For example, if you wanted to allow people from all over the world to query some database you had developed, you could create an executable CGI script that would transmit information to the database engine and then receive results and display them in the user's Web browser. The user could not directly access the database without some gateway to allow access. This link between the database and the user is the "gateway," which is where the CGI standard originated.

A CGI script can be written in any language that allows it to be executed (e.g., C/C++, Fortran, PERL, TCL, Visual Basic, AppleScript, Python, and Unix shells), but by far, the most common language for CGI scripting is PERL, followed by C/C++. A CGI script is easier to debug, modify, and maintain than the typical compiled program, so many people prefer CGI for this reason.

For tutorials on CGI, please see http://cgi.resourceindex.com/Documentation/.

Significance

The importance of CGI lies in the fact that its flexibility has made it a standard for running executable files from Web servers. This standard allows for true interactivity, in endless ways, on Web sites. For example, CGI scripts can implement the following kinds of features (http://cgi.resourceindex.com/Programs_and_Scripts/, http://www.cpan.org/ http://worldwidemart.com/scripts/, and http://bignosebird.com/cgi.shtml provide thousands of free, "canned" scripts for these uses):

Access Counters	Display the number of visitors to your site in a text-based or graphical manner
Advertisements	Set up banner rotations on your Web page and track their statistics

Auctions	Provide for Web-based auctions
Audio Management	Provide and manage audio files in different formats for users to listen to
Bulletin Board Message Systems	Provide on-line message forums for threaded discussions
Calendars	Schedule events and/or allow users to post dated information
Chat	Provide for real-time chats on the Web
Classified Ads	Allow users to post information on buying, selling, and/or trading possessions
Clocks	Display the current time on your Web page in an image- or text-based format
Commerce and Finance	Allow users to use calculators, credit cards, etc.
Content Retrieval	Retrieve and integrate content of all kinds into your Web site (e.g., news and headlines, stock quotes, weather)
Cookies	Track visitors and store user information
Countdowns	Display on your Web page the length of time until a specified event
Customer Support	Maintain knowledge bases; provide for customer support e-mail ticketing, real-time customer support, and FAQ maintenance
Database Manipulation	Create, edit, and manipulate databases; allow users to search your databases
Development Environments	Aid in the development of programs and promote collaboration
Editing Web Pages	Allow users or administrators to create and edit Web pages

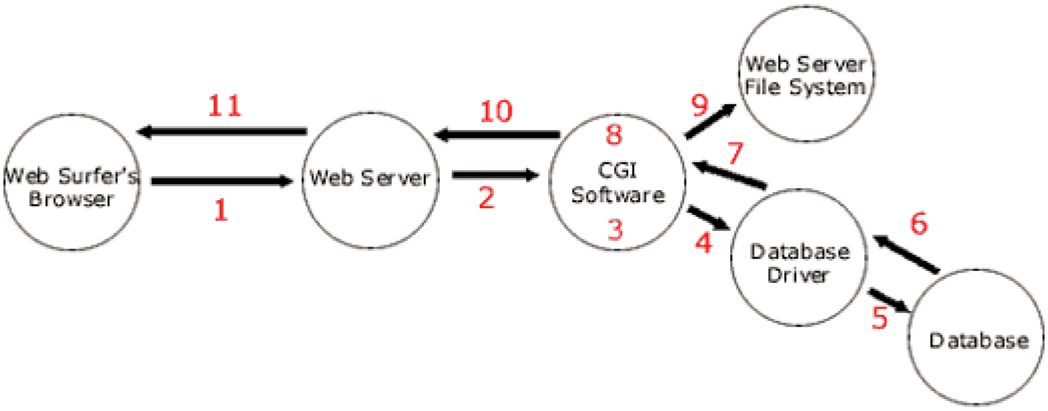
File Management	Manage your files and directories via the Web (e.g., file downloading/uploading, link protection)
Form Processing	Process basic forms and send results via e-mail
Games	Allow users to play games on your Web site
Guestbooks	Allow visitors to sign in and leave a message on your Web site
HTML Manipulation	Create and insert tables, frames, lists, headers, and footers
Homepage Communities	Allow visitors to create their own customized Web pages on your site
Image Display	Display images in different ways and formats; index images; allow for picture posting; timed rotation of images
Imagemaps	Create clickable images that redirect the user to another page
Instant Messaging	Allow users to participate in instant messaging or to use various features of instant messaging systems
Interactive Stories	Allow visitors to add to an existing story
Internet Utilities	Provide for a wide range of standard Internet programs (e.g., finger, telnet, traceroute, whois)
Link Indexing Scripts	Allow visitors to add links to your Web site
Link Verification	Test the links on your page; evaluate your server and its performance
Logging Accesses and Statistics	Track the number of visitors to your site; log useful statistics regarding your site and visitors

Mailing Lists	Allow for mailing lists and management of existing mailing lists
News Posting	Post and manage news and updates to your site
Password Protection	Password protect your Web site
Postcards	Allow visitors to your Web page to send an Internet postcard to someone
Random Items	Include random links, text, images, pages, etc. on your Web site
Redirection	Redirect users in various ways (e.g., jump boxes, browser-based, error-based)
Reservations and Scheduling	Make reservations and usage schedules for items, people, etc.
Searching	Allow users to perform searches on your Web site using keywords and phrases
Shopping Carts	Set up an on-line store using shopping cart and catalog features
Spam Prevention	Randomly generate fake (but real-looking) e-mail addresses for spammer robots to pick up, resulting in a bunch of bounced spam for the spammer and less spam for you and your users
Surveys and Voting	Conduct surveys, take votes, allow users to post ratings and reviews
Tests and Quizzes	Create and grade tests and quizzes automatically over the Web
Web Server Maintenance	Maintain and monitor your server
Web-Based E-Mail	Supply e-mail access to users through their Web browser

Website Promotion	Set up affiliate programs for sites referring visitors to you; set up contests/awards to attract visitors to your site; set up a link exchange between sites with banner ads; allow users to recommend your Web site to a friend; submit your URL to search engines; track referrers; create your own ring of Web sites

How CGI Scripts Work

Http://www.linkyours.com/cgi_overview.html provides an excellent overview (reproduced here with permission) of the process by which CGI scripts are executed:



1. The Web surfer fills out a form and clicks, “Submit.” The information in the form is sent over the Internet to the Web server.
2. The Web server “grabs” the information from the form and passes it to the CGI software.
3. The CGI software performs whatever validation of this information that is required. For instance, it might check to see if an e-mail address is valid. If this is a database program, the CGI software prepares a database statement to either add, edit, or delete information from the database.
4. The CGI software then executes the prepared database statement, which is passed to the database driver.
5. The database driver acts as a middleman and performs the requested action on the database itself.
6. The results of the database action are then passed back to the database driver.
7. The database driver sends the information from the database to the CGI software.
8. The CGI software takes the information from the database and manipulates it into the format that is desired.
9. If any static HTML pages need to be created, the CGI program accesses the Web server computer’s file system and reads, writes, and/or edits files.

10. The CGI software then sends the result it wants the Web surfer's browser to see back to the Web server.
11. The Web server sends the result it got from the CGI software back to the Web surfer's browser.

Implementing CGI Scripts

In general, there are several steps necessary for installing CGI scripts that are the same regardless of the kind of script you are trying to install. Since I am assuming that the reader is a novice, the following instructions describe only how to *install* pre-made CGIs. For more information on how to *write* scripts, please refer to the CGI tutorial listed above or to one of the sites listed in the References section. (Also, for a more detailed explanation of the installation process described below, please see http://adashimar.hypermart.net.)

First, you should know how to use a text editor (e.g., Notepad). CGI scripts are no different from static text files; they only become dynamic with a CGI file extension (e.g., .pl or .cgi). Both are written in ASCII (e.g., text) and can be opened and manipulated using any text editor.

You must also know the path to the compiler on your host server. For instance, if you are using Perl, the path to the Perl compiler on your Web host would look something like "#!/usr/bin/perl" or "#!/usr/local/bin/perl". This path will always be the first line in your script. If you do not know the path to the compiler on your server, you can find out by using telnet to get to your server and then using a "whereis" command (e.g., "whereis perl), or by simply asking your host provider for the information. Similarly, you must know the path to your site from your server. It should look something like this: "/usr/local/etc/httpd/sites/mysite.com". Depending on the type of CGI script you are trying to implement, you may also need to know the path to your server's mail program or your cgi-bin. You can get this information by checking the CGI help section of your host provider's Web site or by asking them directly. Once you have obtained all the necessary path information, you can simply use a text editor to change the CGI script to reflect your own path information.

Finally, you will need to upload the script to your server. Conceptually, this means using an FTP program to upload the script to your cgi-bin directory. The default on your FTP program, however, will be to upload files in *binary* mode. While this is acceptable for most files, CGI files *must* be uploaded in ASCII format. Otherwise, they will not work. Remember, then, to always switch to ASCII mode before uploading CGI scripts.

After uploading the CGI script, it is important to set permission for it (see the discussion of security issues for CGI below). This tells your server who can read, write (i.e., modify), and execute your script. First check the documentation inside the script to find out whether you should "chmod 755" or "chmod 777" it. "Chmod" refers to the "change mode" command, and the numbers simply designate two different types of permission settings (for a more thorough discussion of these settings, please refer to http://jgo.local.net/LinuxGuide/linux-chmod.html). Once you have determined the proper permission setting for your script, permission can be set one of three ways: (1) using your FTP program, (2) by telnetting to your server and using the "chmod" command, or (3) by asking the support personnel at your Web host to change permission settings for you. If you wanted to change permission settings using the FTP program, WSFTP, for instance, you could do this by right-clicking on the script and selecting "chmod." This will bring up a "Remote file permissions" box with three users for whom you must set permission (i.e., owner, group, and other) and three permission levels (i.e., read, write, and execute). If your script designates a chmod 755 permission level, you will allow the owner all three permission levels, and allow "group" and "other" read and execute permissions, but not write permission. Basically, this means that you are allowing users a read-only version of your file. They will be allowed to execute the script, but not to change it. If your script designates a chmod 777 permission level, you allow all users all levels of permission. Http://jgo.local.net/LinuxGuide/linux-chmod.html, however, warns of the dangers with allowing chmod 777 permission levels on a CGI script, as this basically "allows the world to replace the program with whatever they'd like."

An Example

The following CGI script was created using Perl. After users vote for their favorite names (another CGI script), this script tabulates votes to create the table of output following the script:

```
#!/usr/local/bin/perl

print "Content-type:text/html\n\n";

open(INF,"votes.out");
@NAMES = <INF>;
close(INF);

foreach $line (@NAMES)
{
   $linecount++;
   @values = split(/\|/, $line);
   foreach $value (@values)
   {
         $FORM{$value}++;
   }
}

print "<html><head><title>Current Results</title></head>\n";
print "<BODY BGCOLOR=\"#AABBBB\" TEXT=BLACK LINK=\"#001170\"
VLINK=\"#001170\" ALINK=\"#001170\">\n";
print "<h2>Current Results</h2><BR><BR>\n";
print "<TABLE WIDTH=400 BORDER=1 CELLSPACING=0 CELLPADDING=0
BGCOLOR=\"#779E9E\">";
print "<TR><TD colspan=1 align=center><h3>Boy Names:</h3></TD>";
print "<TD colspan=1 align=center><h3>Girl Names:</h3></TD>";
print "<TR BGCOLOR=\"#88AFAF\"><TD><TABLE WIDTH=200 BORDER=0
CELLSPACING=0 CELLPADDING=0";
print " BGCOLOR=\"#779E9E\">\n";
@boynames = ("boyname a","boyname b","boyname c");

foreach $x (@boynames) {
   {
         print "<TR BGCOLOR=\"#88AFAF\"><TD WIDTH=33%>&nbsp;</TD>";
      print "<TD WIDTH=30%>$x</TD><TD
WIDTH=4%><B>$FORM{$x}&nbsp;</B></TD>";
      print "<TD WIDTH=33%>&nbsp;</TD></TR>\n";
   }
}
print "</TABLE></TD><TD>";
print "<TABLE WIDTH=200 BORDER=0 CELLSPACING=0 CELLPADDING=0";
print " BGCOLOR=\"#779E9E\">\n";

@girlnames = ("girlname a", "girlname b", "girlname c");

foreach $x (@girlnames) {
   {
         print "<TR BGCOLOR=\"#88AFAF\"><TD WIDTH=33%>&nbsp;</TD>";
         print "<TD WIDTH=30%>$x</TD><TD
WIDTH=4%><B>$FORM{$x}&nbsp;</B></TD>";
         print "<TD WIDTH=33%>&nbsp;</TD></TR>\n";
   }
}
print "</TABLE></TD></TR>";
print "<TR><TD COLSPAN=2 BGCOLOR=\"#779E9E\"
ALIGN=RIGHT><B>$linecount</B> people have voted.</TD></TR>";
print "</TABLE>";
print "<BR><BR><BR>\n";
print "<A HREF=\"http://www.beth.cx/baby/\">Go back to Beth v2.0</A>\n";
print "</DIV>\n";
print "</BODY></HTML>\n";
```

Current Results

Boy Names:				Girl Names:			
	boyname a	**6**			girlname a	**6**	
	boyname b	**11**			girlname b	**7**	
	boyname c	**1**			girlname c	**7**	
20 people have voted.							

Why CGI?

The above applications can be implemented using other means as well (e.g., server-side JavaScript, PHP, ACGI, VRML, DHTML), but many of these other means developed after CGI. CGI, then, has become a standard, and many programmers prefer simply to "tweak" their old CGI scripts for new purposes, instead of starting from scratch with the newer languages. Also, CGI is more versatile in many ways. A traditional CGI application using Perl, for instance, can be run on a large number of platforms with a wide variety of Web servers. A programmer using server-side JavaScript, however, would be limited to Netscape Enterprise Server.

CGI has its disadvantages though. Many of the newer languages developed in response to CGI being slow, so they are significantly faster. An informal study of the performance of CGI versus server-side JavaScript, for instance, found the following discrepancy in access times for two Web pages:

	Category page	**Product page**
CGI/Perl	219 ms	5990 ms
LiveWire/JavaScript	198 ms	104 ms

Also, there are significant security issues with CGI. Since a file that uses CGI is executable, it is equivalent to letting anyone in the world run a program on your machine. Obviously, this is not the safest thing to do. For this reason, many Web hosts do not allow users to run CGI scripts. In this case, though, you can have your CGI applications hosted for you remotely. Http://www.hypermart.net is an almost-free host that allows CGI scripting, and http://cgi.resourceindex.com/Remotely_Hosted/ lists a number of other hosts that allow CGI.

Related is the fact that programs that use CGI scripts need to reside in a special directory, so that the server knows to execute the program rather than simply display it to the browser. This directory, commonly /cgi-bin, is under the direct control of the Webmaster. This prohibits the average user from creating and running programs that use CGI.

For more information on CGI and security issues, please see http://www.w3.org/Security/Faq/wwwsf4.html.

Summary

CGI is a standard for interfacing executable files with Web servers. It allows for the interactive, dynamic, flexible features that have become standard on many Web sites, such as guestbooks, counters, bulletin boards, chats, mailing lists, searches, shopping carts, surveys, and quizzes. Several newer, faster means for accomplishing these same kinds of tasks have been developed, but CGI is more flexible in a number of ways. CGI is commonly used whenever one needs a Web server to run a program in real-time, take some kind of action, and then send the results back to a user's browser. Scripts can be written in any language that allows a file to be executed, but the most common language for CGI scripts is Perl. One does not have to be a programmer to use CGI scripts (although this helps!), as there are a number of sites that offer free, "canned" scripts that can be modified with the installer's personal server path information.

References

http://adashimar.hypermart.net/ -- introduction to CGI; focuses on steps for installing

http://bignosebird.com/cgi.shtml -- free Perl CGI scripts

http://cgi.resourceindex.com -- CGI-related resources, including scripts (/Programs_and_Scripts/), documentation (/Documentation), and information on remote hosting for CGI (/Remotely_Hosted/)

http://developer.netscape.com/viewsource/lazar_cgi.html -- CGI vs. server-side JavaScript for databases

http://hoohoo.ncsa.uiuc.edu/cgi/overview.html -- CGI documentation (including a good intro page (/intro.html) and examples

http://hotwired.lycos.com/webmonkey/99/26/index4a_page4.html?tw=programming -- CGI permission levels

http://jgo.local.net/LinuxGuide/linux-chmod.html -- CGI permission levels and chmod

http://worldwidemart.com/scripts/ -- free CGI scripts

http://www.cpan.org/ -- Comprehensive Perl Archive Network; documentation, FAQ, mailing list, scripts

http://www.hypermart.net -- Web host that allows CGI scripting http://www.linkyours.com/cgi_overview.html -- graphic overview of CGI interface process

http://www.mattkruse.com/info/cgi/ -- good, simple introduction to CGI

http://www.w3.org/Security/Faq/wwwsf4.html -- World Wide Web Security FAQ on CGI scripting

This paper was written by Alicia D. Beth for the course EDC385G,
Multimedia Authoring, at the University of Texas - Austin